J. Aust. Math. Soc. **85** (2008), 197–209
doi:10.1017/S1446788708000864

GROUPS AND SEMIGROUPS WITH A ONE-COUNTER WORD PROBLEM

DEREK F. HOLT✉, MATTHEW D. OWENS and RICHARD M. THOMAS

(Received 29 December 2007; accepted 21 May 2008)

Communicated by Peter M. Neumann

Dedicated to Cheryl Praeger for her sixtieth birthday

Abstract

We prove that a finitely generated semigroup whose word problem is a one-counter language has a linear growth function. This provides us with a very strong restriction on the structure of such a semigroup, which, in particular, yields an elementary proof of a result of Herbst, that a group with a one-counter word problem is virtually cyclic. We prove also that the word problem of a group is an intersection of finitely many one-counter languages if and only if the group is virtually abelian.

2000 *Mathematics subject classification*: primary 20F10, 20M45; secondary 68Q45, 03D40.
Keywords and phrases: groups, semigroups, word problem.

1. Introduction

There are several intriguing connections between group theory and formal language theory. For example, we can consider groups G whose word problem lies in a particular class $\mathcal{F}$ of languages. We take a finite group generating set X for G and let A be the disjoint union of X and X^{-1}; we then have a natural (monoid) homomorphism φ from A^* onto G (where A^* represents the set of all finite words in the symbols A, including the empty word ε). We define the *word problem* of G to be $1\varphi^{-1}$; that is, the set of words in A^* that represent (via φ) the identity in G.

It would appear that whether or not the word problem lies in the family $\mathcal{F}$ depends on the choice of X, but it is well known that this is not the case if $\mathcal{F}$ is closed under inverse homomorphism (see [9] for example).

Observe that considering words representing the identity is sufficient to decide whether two words u and v in A^* represent the same element of G, since this is the case if and only if uV represents the identity (where V is the word obtained from v by replacing each symbol by the corresponding inverse symbol and then reversing the word). We can think of the word problem of the group as being the set of all words uV such that u and v are words in A^* representing the same element of G.

An obvious question is the extent to which this generalizes to semigroups. In [3] Duncan and Gilman take the following definition for the word problem of a semigroup S. If A is a set of semigroup generators for S (so that each element of S can be represented by a word in A^+, the set of all nonempty words over A), then the word problem of S with respect to A is

$$\{u\#v^{\mathrm{rev}} \in (A \cup \{\#\})^+ \mid u, v \in A^+, u =_S v\}.$$

Here # is a symbol not in A, v^{rev} denotes the reversal of the word v, and $u =_S v$ means that the words u and v represent the same element of S. If we want to stress that u and v are identical as strings, we write $u \equiv v$.

Given this, we can talk about the word problem of a semigroup lying in a class $\mathcal{F}$ of languages. As with groups, if $\mathcal{F}$ is closed under inverse homomorphism, then membership of the word problem in $\mathcal{F}$ is independent of the choice of finite generating set for the semigroup. The definition in [3] is a natural extension of the notion of the word problem from groups to semigroups since the word problem of a group in the group sense lies in $\mathcal{F}$ if and only if the word problem in the semigroup sense lies in $\mathcal{F}$.

In this paper we will concentrate on semigroups whose word problem is a one-counter language; the one-counter languages form an interesting class lying strictly between the classes of regular and context-free languages. We define the one-counter languages and mention some of their properties in Section 2.

In the group case the one-counter languages were shown in [8] to be particularly significant in the following sense. If $\mathcal{F}$ is a class of languages that forms a 'cone' (that is, if $\mathcal{F}$ is closed under homomorphism, inverse homomorphism and intersection with regular languages) and if $\mathcal{F}$ is contained in the class of context-free languages, then the class of groups whose word problem lies in $\mathcal{F}$ coincides with the groups with regular, one-counter or context-free word problem.

It is well known [1] and straightforward to prove that the groups with regular word problems are precisely the finite groups. Duncan and Gilman comment in [3] that this generalizes to semigroups and, for completeness, we include a proof of this fact here (see Proposition 3.1). As far as context-free languages are concerned, there is the lovely characterization by Muller and Schupp [12] to the effect that a group has a context-free word problem if and only if it is virtually free (that is, if it has a free subgroup of finite index). This is a deep result and uses, amongst other things, Stallings' characterization [15] of groups with more than one end. In addition, the result in [12] that a context-free group is virtually free requires an extra hypothesis, that of 'accessibility', and the need for this was only removed by another deep result, namely that any finitely presented group is accessible (which was proved by Dunwoody in [4]). As a result of all this, Muller and Schupp's proof is both ingenious in its ideas and also uses some deep results.

Given these results, in the group case this leaves the one-counter groups. Herbst [8] showed that a group has a one-counter word problem if and only if it is virtually cyclic. Not surprisingly, he uses the characterization by Muller and Schupp and so, as it stands, a complete proof of his result also requires some complex group theoretic

machinery. In our investigation of semigroups with a one-counter word problem, we prove a result (see Theorem 4.2) that gives a very strong condition on the structure of such a semigroup. One consequence is that we can deduce Herbst's result from this and hence give an elementary proof of his result (see Theorem 4.3).

We conclude the paper by proving in Theorem 5.2 that a group is virtually abelian of free rank at most n if and only if its word problem is an intersection of n one-counter languages. This is related to a recent result of Elder, Kambites and Ostheimer [5] that a group is virtually abelian if and only if its word problem is recognized by a G-automaton with G free abelian. However, neither of these two results is a direct consequence of the other in any obvious way. In both cases, the proofs make use of a deep result of Gromov, that a finitely generated group has polynomial growth function if and only if it is virtually nilpotent [7]. Note, however, that our elementary proof of Theorem 4.3 does not use Gromov's result.

2. One-counter languages

In this section we summarize some properties of the one-counter languages that will be relevant for this paper; our main source of information here is [2] and the reader is referred there for further details.

It is well known that the context-free languages are accepted by pushdown automata; however, there are some variations in convention when defining pushdown automata (although these make no difference to the class of languages accepted). In this paper, a *pushdown automaton* M will be a sextuple $(Q, \Sigma, \Gamma, \tau, s, A)$, where Q, Σ and Γ are finite sets (Q is the set of 'states', Σ the 'input symbols' and Γ the 'stack symbols'); there is a special 'bottom of stack symbol' $\perp$ in Γ. In addition, the transition relation τ is a finite subset of $Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \times Q \times \Gamma^*$, $s \in Q$ and $A \subseteq Q$, and we insist that

$$\begin{aligned} (q, a, \perp, r, \gamma) \in \tau \quad &\Rightarrow \quad \gamma \in \{\perp\}\Gamma_1^*, \\ (q, a, g, r, \gamma) \in \tau, g \in \Gamma_1 \quad &\Rightarrow \quad \gamma \in \Gamma_1^*, \end{aligned}$$

where $\Gamma_1 = \Gamma - \{\perp\}$. In other words, $\perp$ appears on the bottom of the stack and nowhere else. There are different notions of acceptance. One possibility is to accept by accept state so that, if a computation is started with the machine in state s with only $\perp$ on the stack, then the input is accepted if it is possible to be in a state in A at the end of the computation. Another possibility is to accept by empty stack (that is, the stack must be empty once the input has been read). We can also insist on both happening simultaneously (that is, a word is accepted if the machine is in an accept state with an empty stack at the end of the computation). These are all equivalent in the sense that, if a language is accepted by a machine with one notion of acceptance and we choose a different notion, then there is a machine with the second notion of acceptance accepting precisely the same language.

If there is no stack, so that we just have a quintuple (Q, Σ, τ, s, A), then we have a finite automaton, and finite automata accept the class of regular languages.

Given this definition of a pushdown automaton, we define a *one-counter automaton* to be a pushdown automaton $(Q, \Sigma, \Gamma, \tau, s, A)$ with $|\Gamma| = 2$. If $\Gamma = \{\perp, g\}$ then, at any stage, the stack of the machine contains $\perp g^n$ for some $n \geq 0$, and so is effectively described by a single natural number n; hence the title 'one-counter'. If, in addition, the pushdown automaton is deterministic (that is, if, for any configuration of the machine, there is at most one possible transition), then we say that we have a *deterministic one-counter automaton*. We say that L is a *one-counter language* if $L = L(M)$ for some one-counter automaton M, and a *deterministic one-counter language* if $L = L(N)$ for some deterministic one-counter automaton N. As with the context-free languages, insisting on determinism does restrict the class of languages accepted.

In general, the closure properties of a class of languages are very important; this is particularly true when considering word problems of groups and semigroups. The class of one-counter languages is closed under each of the following operations (see [2]):

- union, concatenation, intersection with regular sets;
- Kleene star;
- gsm mappings, inverse gsm mappings.

(A gsm mapping is one defined by a finite state transducer.) As a homomorphism is a special case of a gsm mapping, the class is closed under homomorphisms and inverse homomorphisms. We will use these closure properties without further comment in what follows.

The deterministic one-counter languages admit some of these closure properties (for example, they are closed under inverse gsm mappings and intersection with regular sets) but not all of them (for example, they are not closed under union).

3. Word problems

As we mentioned in Section 1, Anisimov showed [1] that a group has a regular word problem if and only if it is finite. It is mentioned in [3] that this generalizes to semigroups; for completeness, we include a proof here.

PROPOSITION 3.1. *A semigroup S has a regular word problem if and only if S is finite.*

PROOF. Let S be a semigroup with a finite generating set A.

Suppose that S is finite of order n. We construct a finite state automaton that consists of the Cayley graph of S together with an additional start state, with arrows labelled a from the start state to the vertex representing a for each $a \in A$. By choosing the vertex representing s as the unique accepting state, we see that the set of words in A^+ representing each element s of S is a regular language L_s. Then the word problem of S is the union over $s \in S$ of the regular languages $L_s \# L_s^{\text{rev}}$ and is therefore regular.

Conversely suppose that the word problem W is regular. For each word $u \in A^+$, let $\sigma(u)$ be the state of a deterministic finite state automaton recognizing W immediately

after reading u in an accepting path for $u\#u^{\text{rev}}$. Then, if u and v represent distinct elements of S, we must have $\sigma(u) \neq \sigma(v)$ since otherwise the automaton would accept $u\#v^{\text{rev}}$. Since there are only finitely many possible $\sigma(u)$, S must be finite. □

The following result is well known for groups; see [10] for example. It is easily generalized to semigroups.

PROPOSITION 3.2. *Let $\mathcal{F}$ be a class of languages closed under inverse homomorphisms and intersection with regular sets; then the class of semigroups whose word problem lies in $\mathcal{F}$ is closed under taking finitely generated subsemigroups.*

When H is a subgroup of finite index in a group G, then we call G a *finite index overgroup* of H. The following result from [10] will be useful.

PROPOSITION 3.3. *Let $\mathcal{F}$ be a class of languages closed under union with regular sets and inverse gsm mappings; then the class of $\mathcal{F}$-groups is closed under passing to finite index overgroups.*

In particular, Propositions 3.2 and 3.3 apply when $\mathcal{F}$ is the class of regular, deterministic one-counter, one-counter, deterministic context-free or context-free languages. It also applies if $\mathcal{F}$ is the family of languages obtained by taking finite intersections of languages from any one of these classes (this yields nothing new in the case of the regular languages, as they are closed under intersection, but does yield new families of languages in the other four cases).

4. One-counter groups and semigroups

For a semigroup S with specified finite generating set A and $n \in \mathbb{N}$, let $\gamma_{S,A}(n)$ be the number of elements of S that are represented by words in A^+ of length at most n. Then $\gamma_{S,A}$ is called the *growth function* of S with respect to A. Properties such as whether the growth function is linear, polynomial of degree d, or exponential are generally independent of the chosen finite generating set for S. (Note also that some authors use the number of elements of S represented by words of length exactly n. If we were using that convention, then, by linear growth, we would mean that this number was uniformly bounded. So, for example, whichever convention we are using, a free abelian group of rank one has linear growth, but one of rank two or more does not.) Our first result about semigroups with a one-counter word problem is the following proposition.

PROPOSITION 4.1. *If a finitely generated semigroup S has word problem a one-counter language, then S has a linear growth function.*

PROOF. Let A be a finite set that generates S and let q be the number of states of a one-counter automaton M accepting the word problem of S with respect to A. We assume that M accepts by empty stack.

For each word $w \in A^+$, choose a shortest computation path $p(w)$ in M that accepts $w\#w^{\text{rev}}$. Let $|w|$ denote the length of w. We shall show that there is a constant K,

which does not depend on w, such that, immediately after reading the symbol # in $p(w)$, the stack height $h(w)$ is at most $K|w|$. If we can show this then, for words w of length at most n, there are only $(Kn+1)q$ possibilities for the pair $(h(w), t(w))$, where $t(w)$ is the state of the machine immediately after reading the symbol # in $p(w)$. This pair cannot be the same for two words w_1 and w_2 that represent different group elements, because if they were then there would be an accepting path for $w_1 \# w_2^{\text{rev}}$. Hence, given this, the growth function $\gamma_{S,A}$ satisfies $\gamma_{S,A}(n) \leq (Kn+1)q$, and the result follows.

We need to prove the claim that $h(w) \leq K|w|$ for some $K \geq 0$. We can assume, without loss of generality, that all moves in M change the stack height by at most one. Moves are either reading moves, when one input symbol is read, or else nonreading moves. We will assume also that the stack height is not changed by reading moves. (We can achieve this by breaking up a reading move that alters the stack into two moves.)

Let w be a word with $|w| = n$. We shall refer to the initial part of $p(w)$, up to and including the reading move in which we read #, as the part in which we read w, and to the remainder of $p(w)$ as the part in which we read w^{rev}. So, with this convention, $h(w)$ is the height of the stack in $p(w)$ after reading w. If $h(w) > q(n+1)$, then, when reading w in $p(w)$, there must be at least one occasion where, between reading two input symbols (or before reading the first symbol), the stack height increases by at least q. Whilst this is happening, M must repeat states. In fact, we can find a subpath of $p(w)$ (which we shall refer to as a *circuit*) linking the repeated states in which the stack height is increased by r for some r with $0 < r \leq q$. Similarly, when reading w^{rev} in $p(w)$, there must be a gap between reading input symbols (or after reading the final symbol), and a circuit linking repeated states in this gap in which the stack height is decreased by u for some u with $0 < u \leq q$. If $h(w) > q^3(n+1)$, then we can find gaps between reading input symbols containing q^2 disjoint circuits of this kind in $p(w)$, in which the stack height is increased by at most q when reading w, and decreased by at most q when reading w^{rev}. Amongst the q^2 circuits in which the height is increased, at least q of them must increase it by the same number $r \leq q$. Similarly, amongst the q^2 circuits in which the height is decreased, at least q of them must decrease it by the same number $u \leq q$.

The idea is to remove u of the circuits that increase the stack height by r, and r of the circuits that decrease the stack height by u, to produce a shorter path accepting $w\#w^{\text{rev}}$, thereby contradicting the minimality of $p(w)$. For this to work, we have to make sure that the stack cannot become empty at any stage between the removed increasing circuits and the removed decreasing circuits, since this would alter the computation.

To do this, we assume that $h(w) > q^3(n+2)$. We choose the gap in which we remove the circuits while reading w to be the latest one in which the stack height increases by at least q^3 at some stage during the gap, and we remove the circuits as late as possible during that gap. Similarly, we choose the gap in which we remove the circuits while reading w^{rev} to be the earliest one in which the stack height decreases by at least q^3 at some stage during the gap, and we remove the circuits as early as

possible during that gap. Between the first of these gaps and the end of w, the stack cannot empty, since otherwise there would certainly be an increase of q^3 in a later gap. Between the beginning of w^{rev} and the place where the decreasing circuits were removed, the stack height decreases by less than q^3 during each gap, and hence by less than $q^3(n+1)$ altogether. Since we are only removing some of the circuits, it could also decrease by some number less than q^3 during the part of $p(w)$ in which the circuits are removed. But, since $h(w) > q^3(n+2)$, the stack can never empty.

This contradiction proves that $h(w) \leq q^3(n+2)$. Thus we have proved the claim, and also the theorem. □

Given Proposition 4.1, we can now prove a result that imposes a very tight restriction on the structure of a semigroup with a one-counter word problem. For the proof we use the idea of the ShortLex ordering $<_{\mathrm{SL}}$ on words, which is defined as follows (as in [6] for example). Suppose that X is a finite set with a linear order $<_X$. Then we say that $\alpha <_{\mathrm{SL}} \beta$ (where $\alpha, \beta \in X^*$) if either:

(1) $|\alpha| < |\beta|$; or
(2) $\alpha \equiv a_1 a_2 \cdots a_m$, $\beta \equiv b_1 b_2 \cdots b_m$, and there exists k with $1 \leq k < m$ such that $a_1 = b_1, a_2 = b_2, \ldots, a_{k-1} = b_{k-1}, a_k <_X b_k$.

The ordering $<_{\mathrm{SL}}$ is a well-ordering on X^*. We shall call the least representative of a semigroup element under $<_{\mathrm{SL}}$ the ShortLex normal form of that element. Note that, if uvw is a word in ShortLex normal form, then so is v, because $v' <_{\mathrm{SL}} v$ with $v =_S v'$ implies $uv'w <_{\mathrm{SL}} uvw$ with $uv'w =_S uvw$. In other words, all contiguous subwords of a word in ShortLex normal form, including all of its prefixes and suffixes, are themselves in ShortLex normal form.

THEOREM 4.2. *If S is a finitely generated semigroup with linear growth then there exist finitely many elements $a_i, b_i, c_i \in S \cup \{\varepsilon\}$ such that every element of S is represented by a word of the form $a_i b_i^n c_i$ for some i and some $n \geq 0$.*

PROOF. Let S be a semigroup with linear growth generated by a finite set A. Choose a linear order on A and then consider the set N of ShortLex normal forms for S. Let $L \subseteq N$ denote the set of words w in N such that w is a prefix of infinitely many v in N. If S is finite then the result is trivial, so we assume that S is infinite. It is then straightforward to show by induction on k that L contains at least one word of length k for all $k \geq 0$, and so L is infinite. We form a graph Γ with vertex set L and a directed edge from w to wa for each $w \in L$, $a \in A$ such that wa is a word in L. For convenience, we adjoin a 'base-point' as an extra vertex to the graph, with an edge labelled a from the base-point to the vertex labelled a, for each $a \in A$ for which $a \in L$.

Given any vertex, there is a unique path from the base-point to that vertex (since L consists of ShortLex normal forms); thus any vertex has in-degree one, and the graph is a tree rooted at the base-point. For $n \geq 0$, let $K(n)$ be the number of words in L of length n, which is also equal to the number of vertices of Γ at distance n from the base-point. Since Γ is a tree with no finite branches, $K(n)$ is an increasing function of n, but the linear growth of S implies that $K(n)$ is bounded (in fact if $|S|$ has at

most Cn elements represented by words of length at most n, then $K(n) \leq C$ for all n), so there is a constant K with $K(n) = K$ for all sufficiently large n. So, once we are sufficiently far from the base-point, Γ consists of K disjoint paths. Hence Γ is a union of K infinite paths starting at the base-point, which are labelled by infinite words $\sigma_1, \sigma_2, \ldots, \sigma_K$.

Let $w_1, w_2, \ldots, w_K$ be (finite) prefixes of $\sigma_1, \sigma_2, \ldots, \sigma_K$, respectively, such that no w_i is a prefix of another w_j. We shall now define a sequence $p_1, p_2, p_3, \ldots$ of words such that each p_i is equal to one of the words w_j, and $p_1 p_2 \cdots p_i$ is a prefix of σ_1 for all $i \geq 0$. We start by setting $p_1 = w_1$. Then $\sigma_1 = p_1\alpha$ for some infinite word α. But since suffixes of ShortLex normal form words are themselves in ShortLex normal form, α must be equal to one of the words σ_j, and hence α has a prefix p_2 equal to one of the w_j. So $\sigma_1 = p_1 p_2 \beta$ where β has a prefix p_3 equal to some w_j, and so on.

Continuing in this vein, we obtain a sequence $\{p_i\}_1^\infty$ of words from the set $\{w_1, \ldots, w_K\}$. Eventually we must have a repetition of one of the p_i: say $p_i = p_j$ where $i < j$. Then $p_{i+k} = p_{j+k}$ for all $k \geq 0$, as any w_m leads to a line that uniquely determines its suffix. So σ_1 consists of $p_1 \cdots p_{i-1}$ followed by infinitely many repetitions of $y_1 := p_i \cdots p_{j-1}$. Similarly, for each i, the infinite word σ_i is equal to a prefix followed by infinitely many repetitions of some word y_i. Let

$$B = \{b_j \mid b_j \text{ is a cyclic permutation of } y_i \text{ for some } 1 \leq i \leq K\}.$$

Then all of the words in L are of the form $a_i b_j^n$ where $b_j \in B$ and the a_i are finitely many prefixes of the graph.

Now, adjoin the vertex set $M = N - L$ to Γ, with edges from w to wa if and only if $wa \in M$; again we have a unique path to any given vertex. The resulting graph Γ' consists of Γ together with a number of branches of finite length adjoined to some of the vertices of Γ.

If we can prove that the lengths of these new branches are uniformly bounded, then they will represent a finite set of elements c_k, and then all elements of the semigroup will be represented by words of the form $a_i b_j^n c_k$, which will prove the theorem. (This is not quite correct, because some of the new branches may start at vertices representing proper prefixes of the words a_i, but that is not a problem, because there are only finitely many such prefixes.)

If the lengths of the branches are not uniformly bounded, then there exist arbitrarily long branches. So there must exist a prefix u and some $v \in B$ for which there exist arbitrarily long words w such that, for some m (depending on w), $uv^m w \in N$ but $uv^m w_1$ is not infinitely extensible, where w_1 is the first symbol of w. In particular, w does not have v as a prefix.

Suppose the number of semigroup elements of length at most n is at most Cn. As $|w|$ approaches infinity, so must m, so we can choose m and $|w|$ sufficiently large that $m|w| > C(m|v| + |w|)$.

The $m|w|$ subwords of $v^m w$ of the form

$$v^i w(t) \quad \text{for } 1 \le i \le m,\ 1 \le t \le |w|,$$

where $w(t)$ denotes the prefix of w of length t, cannot all represent distinct elements, and since they are all ShortLex geodesics, two of them must be equal as words, $v^i w(s) \equiv v^j w(t)$. Clearly i is not equal to j, as that would give $w(s) \equiv w(t)$ whence $s = t$; thus w has v as a prefix, which is a contradiction and completes the proof. □

We do not know whether the converse of Theorem 4.2 is true, but the following examples show that a finitely generated semigroup with linear growth need not have a one-counter word problem—in fact its word problem need not even be decidable!

Let Z be a subset of $\mathbb{N} - \{0, 1\}$, and let S_Z be the semigroup with zero defined by the presentation

$$S_Z := \langle a, b \mid a^2 = a,\ bab = 0,\ ab^i a = aba\ \forall\, i \in Z \rangle.$$

Then the elements of S_Z are 0, a, b^k $(k > 0)$, ab^k $(k > 0)$, $b^k a$ $(k > 0)$, $ab^k a$ $(k > 0)$, and these are all distinct except that elements $ab^k a$ with $k \in Z$ are equal to aba. Furthermore, each of these words (except for $ab^k a$ with $k \in Z$) is the unique shortest word for the semigroup element that it defines, so S_Z has linear growth.

Now $ab^k a\#aba$ is in the word problem of S_Z if and only if $k \in Z$ and, by intersecting the word problem of S_Z with the language defined by the regular expression $ab^*a\#aba$, we see that this word problem is decidable if and only if Z is recursive.

Theorem 4.2 enables us to provide an elementary proof of Herbst's result. We note that alternative elementary proofs that groups with linear growth are virtually cyclic have been published previously; see for example [11] and [16].

THEOREM 4.3. *A finitely generated group G has a one-counter word problem if and only if G is virtually cyclic.*

PROOF. It is straightforward to show that virtually cyclic groups have a one-counter word problem. Conversely, suppose that G has a one-counter word problem. By Theorem 4.2, there exist elements

$$a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_n$$

in G (for some n) such that

$$G = \bigcup_{i=1}^{n} \bigcup_{r \in \mathbb{N}} a_i b_i^r c_i.$$

(Our convention is that $\mathbb{N}$ contains 0.) Given this, then

$$G = \bigcup_{i=1}^{n} (a_i \langle b_i \rangle a_i^{-1})\, a_i c_i.$$

In other words, G is a union of finitely many cosets. By [13, Lemma 4.1], at least one of the subgroups $a_i \langle b_i \rangle a_i^{-1}$ has finite index in G; so G is virtually cyclic as required. □

5. Intersections of one-counter languages

In this section we characterize those groups whose word problem is a finite intersection of one-counter languages. It is convenient to prove first the following result.

PROPOSITION 5.1. *A finitely generated virtually abelian group of free abelian rank $n \geq 1$ has a word problem that is the intersection of n deterministic one-counter languages.*

PROOF. Suppose that the group G has a free abelian subgroup A isomorphic to $\mathbb{Z}^n$ of finite index. Take a natural group generating set $X = \{x_1, x_2, \ldots, x_n\}$ for A, and let L_i be the language over $X \cup X^{-1}$ consisting of all words w such that the exponent sum of x_i in w is zero (that is, the number of occurrences of x_i in w equals the number of occurrences of x_i^{-1}). We see that L_i is a deterministic one-counter language (we keep track of the exponent sum of x_i on the stack, indicating in the state whether the current sum is positive or negative, and ignore the other inputs). The word problem of A is the intersection of the L_i, and so is the intersection of n deterministic one-counter languages. By Proposition 3.3, the word problem of G is also the intersection of n deterministic one-counter languages (as G is a finite overgroup of A). □

Our characterization is then as follows.

THEOREM 5.2. *The following are equivalent for a finitely generated group G.*

(i) *The word problem of G is the intersection of n one-counter languages for some $n \geq 1$.*
(ii) *The word problem of G is the intersection of n deterministic one-counter languages for some $n \geq 1$.*
(iii) *The group G is virtually abelian of free abelian rank at most n.*

PROOF. The fact that (ii) implies (i) is clear, and (iii) implies (ii) by Proposition 5.1; it remains to show that (i) implies (iii). The proof uses similar ideas to that of [10, Theorem 12].

So suppose that G has a word problem that is the intersection of n one-counter languages. The proof of Proposition 4.1 shows that the stack height of each of the n one-counter automata after reading the symbol # in a minimal length accepting path for $w\#w^{\mathrm{rev}}$ is bounded by a linear function of $|w|$. As in Proposition 4.1, words w_1 and w_2 representing distinct elements of G cannot lead to the same configuration in each of the automata, since otherwise we could construct accepting paths for $w_1\#w_2^{\mathrm{rev}}$ in each automaton. It follows that the growth function of G is polynomial of degree at most n; so, by [7], G is virtually nilpotent. By Theorem 5.3 below, G is virtually abelian, and the result now follows, since a virtually abelian group of free abelian rank n has growth function a polynomial of degree n. □

THEOREM 5.3. *Let G be a finitely generated virtually nilpotent group, and suppose that its word problem is an intersection of finitely many context-free languages. Then G is virtually abelian.*

PROOF. We showed in [10, Theorem 12] that a virtually nilpotent group that is not virtually abelian contains a copy of the Heisenberg group

$$H = \langle A, B, C \mid [A, B] = C, [A, C] = [B, C] = 1 \rangle,$$

where $[x, y]$ denotes the commutator $x^{-1}y^{-1}xy$. As we observed earlier, the property of a word problem being an intersection of finitely many context-free languages is inherited by subgroups and is independent of the generating set. So, in order to prove the theorem, it is sufficient to show that the word problem $W(H, X)$ of H on the generating set $X = \{A, B, C\}$ cannot be an intersection of context-free languages. To do this, we shall show that the intersection W_I of $W(H, X)$ with the regular set $R := (A^{-1})^*(B^{-1})^*A^*B^*(C^{-1})^*$ cannot be an intersection of context-free languages. We observe that, for all $n \geq 0$, $A^{-n}B^{-n}A^nB^nC^{-n^2} \in W_I$, but $A^{-n}B^{-n}A^nB^nC^m \notin W_I$ when $m \neq n^2$.

Suppose, for a contradiction, that $W_I = L_1 \cap \cdots \cap L_m$, with each L_i context-free, where we may assume that each L_i is a subset of R.

In a similar way to that used in [10, Theorem 12], we shall apply Parikh's theorem on bounded context-free languages [14] to the L_i. Recall that, for $n > 0$, a subset L of $\mathbb{N}^n$ is called *linear* if there exist $c \in \mathbb{N}^n$ and a finite subset $P = \{p_1, \ldots, p_k\}$ of $\mathbb{N}^n$ such that

$$L = \left\{ c + \sum_{t=1}^{k} \alpha_t p_t \;\middle|\; \alpha_t \in \mathbb{N} \right\}, \tag{5.1}$$

where we may clearly assume that the p_t are nonzero. Then Parikh's theorem implies that each L_i is a union of finitely many sets $L_{i1}, \ldots, L_{ij_i}$ such that, for each i, j,

$$E_{ij} := \{(a, b, c, d, e) \mid A^{-a}B^{-b}A^cB^dC^{-e} \in L_{ij}\}$$

is a linear subset of $\mathbb{N}^5$. So there are elements $c_{ij}, p_{ij1}, \ldots, p_{ijk_{ij}}$ of $\mathbb{N}^5$ such that (5.1) is satisfied with $L = E_{ij}$.

For elements $v \in \mathbb{N}^5$, we denote the five components of v by $v(1), v(2), v(3), v(4), v(5)$. Let m be the maximum value of $c_{ij}(5)$ for any i, j, and let r be the maximum value of $p_{ijk}(5)/p_{ijk}(l)$ for any i, j, k, l with $1 \leq l \leq 4$ and $p_{ijk}(l) \neq 0$. Let us call the vector p_{ijk} *simple* if its first four components $p_{ijk}(l)$ $(1 \leq l \leq 4)$ are all 0, and *complex* otherwise. Then, if

$$(a, b, c, d, e) = c_{ij} + \sum_{t=1}^{k_{ij}} \alpha_t p_{ijt} \in E_{ij}$$

and if each p_{ijt} for which α_t is nonzero is complex, then we have

$$e \leq m + r(a + b + c + d).$$

Now choose n with $n^2 > m + 4nr$. Since $A^{-n}B^{-n}A^nB^nC^{-n^2} \in W_I$, we have $A^{-n}B^{-n}A^nB^nC^{-n^2} \in L_i$ for $1 \leq i \leq m$ and hence, for each such i, there exists a j with $A^{-n}B^{-n}A^nB^nC^{-n^2} \in L_{ij}$. To ease the notation, let us suppose that $A^{-n}B^{-n}A^nB^nC^{-n^2} \in L_{i1}$ for all i and hence that $(n, n, n, n, n^2) \in E_{i1}$. Since $n^2 > m + 4nr$, it follows from the discussion above that, for each i, the p_{i1t} cannot all be complex and hence there must be a simple p_{i1t}, which we can take to be $p_{i11} = (0, 0, 0, 0, e_i)$ with $e_i > 0$. (We are assuming that the p_{ijk} are nonzero.)

But then $(n, n, n, n, n^2 + te_i) \in E_{i1}$ for all i and all $t \in \mathbb{N}$ and hence, letting e be the least common multiple of the e_i, we have $(n, n, n, n, n^2 + e) \in E_{i1}$ for all i, and hence $A^{-n}B^{-n}A^nB^nC^{n^2+e} \in W_I$. This is a contradiction, because $A^{-n}B^{-n}A^nB^nC^{n^2+e}$ is not equal to the identity element of H. □

Acknowledgements

This paper was completed whilst the third author was on study leave from the University of Leicester and the support of the University in this regard is much appreciated. The third author would also like to thank Hilary Craig for all her help and encouragement. The authors thank the referee for a careful reading of the paper and for raising the question of whether a semigroup satisfying the conclusion of Theorem 4.2 must have a one-counter word problem.

References

[1] A. V. Anisimov, 'Certain algorithmic problems for groups and context-free languages', *Kibernetica* **2** (1972), 4–11.
[2] Jean Berstel, *Transductions and Context-Free Languages* (Teubner, Stuttgart, 1979).
[3] Andrew Duncan and Robert H. Gilman, 'Word hyperbolic semigroups', *Math. Proc. Cambridge Philos. Soc.* **136** (2004), 513–524.
[4] M. J. Dunwoody, 'The accessibility of finitely presented groups', *Invent. Math.* **81** (1985), 449–457.
[5] Murray Elder, Mark Kambites and Gretchen Ostheimer, On groups and counter automata, arXiv:math/0611188v1, 7 Nov 2006.
[6] David B. A. Epstein, J. W. Cannon, D. F. Holt, S. Levy, M. S. Paterson and W. P. Thurston, *Word Processing in Groups* (Jones and Bartlett, Boston, MA, 1992).
[7] Mikhael Gromov, 'Groups of polynomial growth and expanding maps', *Publ. Math. Inst. Hautes Études Sci.* **53** (1981), 53–78.
[8] Thomas Herbst, 'On a subclass of context-free groups', *RAIRO Inform. Théor. Appl.* **25** (1991), 255–272.
[9] Thomas Herbst and Richard M. Thomas, 'Group presentations, formal languages and characterizations of one-counter groups', *Theoret. Comput. Sci.* **112** (1993), 187–213.
[10] Derek F. Holt, Claas E. Röver, Sarah Rees and Richard M. Thomas, 'Groups with a context-free co-word problem', *J. London Math. Soc.* **71** (2005), 643–657.
[11] J. Justin, 'Groupes et semi-groupes à croissance linéaire', *C. R. Acad. Sci. Paris Sér. A–B* **273** (1971), 212–214.
[12] David E. Muller and Paul E. Schupp, 'Groups, the theory of ends, and context-free languages', *J. Comput. System Sci.* **26** (1983), 295–310.
[13] B. H. Neumann, 'Groups covered by permutable subsets', *J. London Math. Soc.* **29** (1954), 236–248.

[14] Rohit J. Parikh, 'Language-generating devices', Quarterly Research Reports, 60, Research Laboratory of Electronics, MIT, Cambridge, MA, 1961, pp. 199–212.
[15] John Stallings, *Group Theory and Three-Dimensional Manifolds*, Yale Mathematical Monographs, 4 (Yale University Press, New Haven, CT and London, 1971).
[16] A. J. Wilkie and L. van den Dries, 'An effective bound for groups of linear growth', *Arch. Math. (Basel)* **42** (1984), 391–396.

DEREK F. HOLT, Mathematics Institute, University of Warwick,
Coventry CV4 7AL, UK
e-mail: D.F.Holt@warwick.ac.uk

MATTHEW D. OWENS, Mathematics Institute, University of Warwick,
Coventry CV4 7AL, UK

RICHARD M. THOMAS, Department of Computer Science, University of Leicester,
Leicester LE1 7RH, UK
e-mail: rmt@mcs.le.ac.uk